DIRECTED VORONOI SEARCH: A direct search method for bound constrained global optimization

Associate Professor Marco Reale
Co-authors: B. L. Robertson, C. J. Price and J. A. Brown

University of Canterbury, Christchurch, New Zealand

December 4, 2013

OVERVIEW

1. Global Optimization
2. Voronoi Partition
3. The DVS optimization algorithm
4. Convergence properties
5. Numerical Results

INTRODUCTION

- The bound constrained global optimization problem may be written as

$$\min f(x) \text{ subject to } x \in \Omega,$$

  where the search region Ω is defined by an n-dimensional box of the form

$$\Omega = \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i \text{ for all } i = 1, \ldots, n\}.$$

- The objective function $f(x)$ maps $\mathbb{R}^n$ into $\mathbb{R} \cup \{+\infty\}$ and is assumed to be lower semi-continuous.

- The objective function can be nonconvex, nonsmooth or even discontinuous.

GLOBAL OPTIMIZATION

- Deterministic Methods
  - Branch and bound, interval analysis, and tunnelling methods.
  - Typically guarantee asymptotic convergence to the global minimum.
- Stochastic Methods (Monte Carlo; random search)
  - Simulated annealing, genetic algorithms, multi-start, clustering algorithms, and partition methods.
  - Ensure convergence to the global minimum in probability.

ACCELERATED RANDOM SEARCH (Stochastic)

- Evaluates the objective function at points selected from a finite sequence of contracting sub-regions (initially Ω itself) centered on the point with the best function value. If a point with a lower function value is found, or if the sequence is exhausted, the search returns back to Ω.

- Accelerated Random Search (ARS) is easy to program, computationally fast, and conceptually simple.

- A downside of ARS is that it can get trapped at a local solution of a multi-modal problem.

ACCELERATED RANDOM SEARCH (Branin test function)

DIRECT (Deterministic)

- Generates a sequence of nested partitions on Ω.
- Each partition consists of a finite number of boxes (initially Ω itself) with the objective function evaluated at the center of each box. At each iteration one or more boxes are divided into three smaller, equally sized boxes using two hyperplanes orthogonal to one of the coordinate axes. The objective function is then evaluated at the center of each new box.
- Dividing a box in this way means that the center point of the old box will be the center point for one of the new boxes.

DIRECT (Branin test function)

VORONOI PARTITION

- A Voronoi partition on Ω using a set of points

$$X = \{x_i \in \Omega : i = 1, 2, \ldots, N\}$$

is defined by the set of N convex polyhedral sub-regions

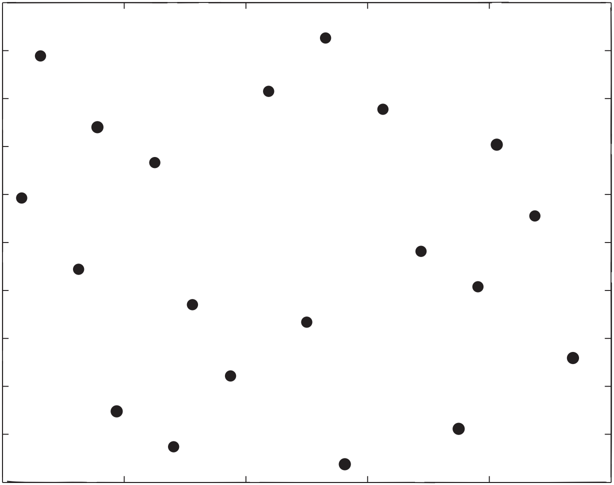
$$A_i = \{x \in \Omega : \|x - x_i\|_2 \leq \|x - x_j\|_2 \text{ for all } j = 1, 2, \ldots, N\},$$

for $1 \leq i \leq N$.

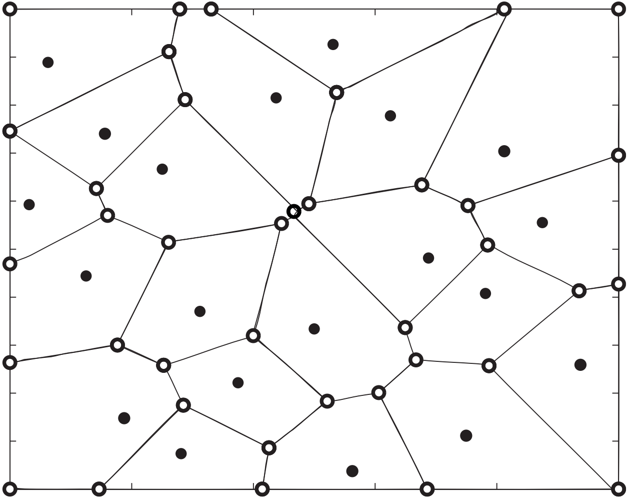
- The Voronoi cell for x_i is A_i, with vertices V_i.

- The set of all vertices, $\mathcal{V} = \{v \in V_i : i = 1, 2, \ldots, N\}$, is called the Voronoi vertex set.

VORONOI PARTITION



VORONOI PARTITION



DVS ALGORITHM (Pseudo Code)

1. **Initialize**: Generate an initial batch of points in Ω and evaluate f at each point. Call this set of points T.
2. **Partition**: Form a Voronoi partition on Ω using T and calculate the Voronoi vertex set V.
3. **Sample**: Select a number of points from V and at least one point from a uniform distribution over Ω. Call this set X and evaluate f at each point in X.
4. **Update T**: Check stopping conditions and if satisfied stop, otherwise $T \leftarrow T \cup X$ and go to step 2.

SELECTION STRATEGIES

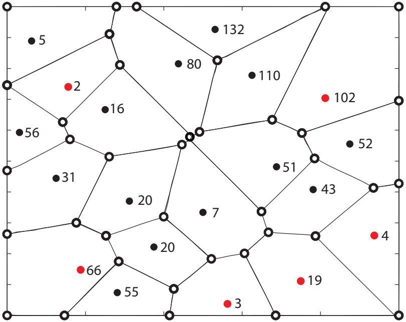
- Choosing points from the Voronoi Vertex set:
  1. Good spread $\rightarrow$ reduce the risk of missing the global minimum.
  2. Increase the point density where f is low $\rightarrow$ increase the rate of local convergence.

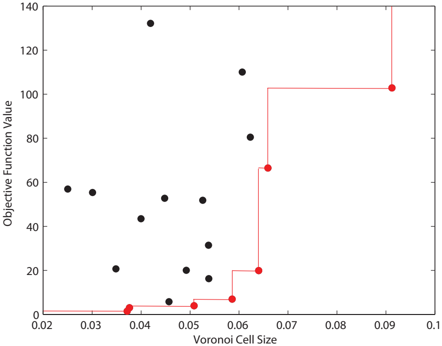
- **(Pareto Optimal Voronoi Cell).** A Voronoi cell containing the sample point x is Pareto optimal unless there exists a cell containing a point z satisfying both

$$f(z) \leq f(x) \quad \text{and} \quad m(z) \geq m(x)$$

  with at least one equality holding strictly, where $m(x)$ is the Lebsegue measure of the cell containing x.

PARETO OPTIMAL VORONOI CELL





- Evaluating f at vertices of Pareto optimal Voronoi cells:
  1. Cells will not remain large.
  2. Point density will increase where f is relatively low.

(I). For each Pareto optimal Voronoi cell V_i, evaluate f at the vertex which has the greatest 2-norm distance from x_i. If $f(x_i) \leq f(x_j)$ for all $x_j \in T$, also evaluate the objective at the vertex which has the least 2-norm distance from x_i.

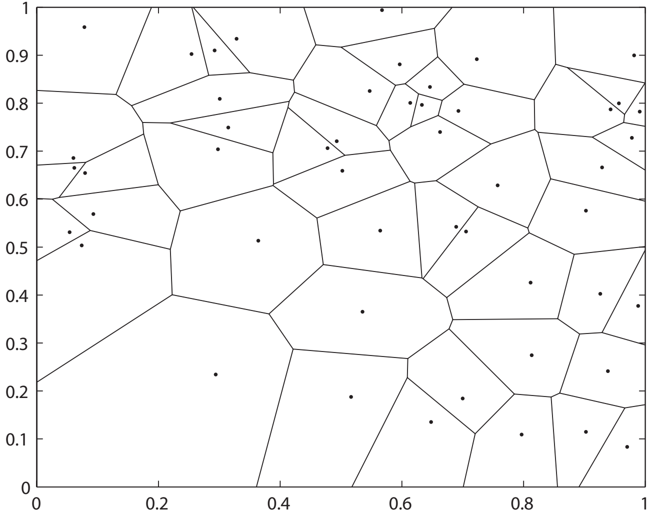
(II). Same as approach (I), but if $f(x_i) \leq f(x_j)$ for all $x_j \in T$, evaluate f at all the vertices in V_i.

- Approach (II) has a stronger local search component, potentially increasing its rate of local convergence.

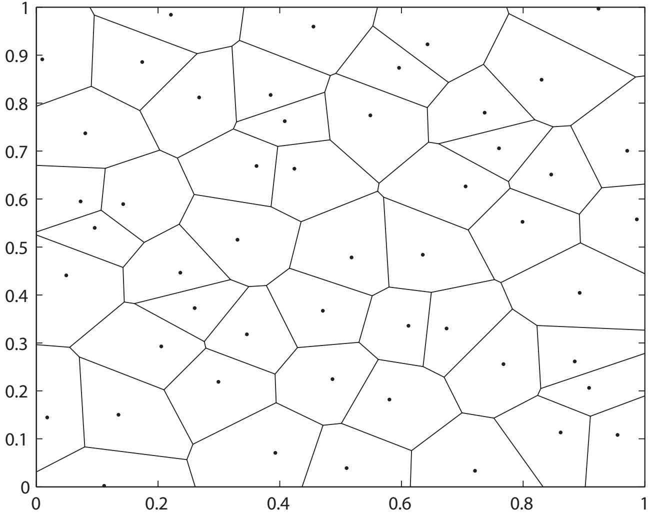
DETERMINISTIC INSTANCE

- Selects points sequentially from the Halton sequence rather than using pseudo-random numbers.
- The ith coordinate of point k may be written in the following form
$$\phi_{b_i}(k) = \sum_{j=0}^{\infty} \left\{ \left\lfloor \frac{k}{b_i^j} \right\rfloor \mod b_i \right\} \frac{1}{b_i^{j+1}},$$
where b_i is the ith prime number.
- The Halton sequence spreads points evenly in low dimensions.

50 PSEUDO-RANDOM POINTS (Poor spread)



50 HALTON POINTS (Good spread)



DETERMINISTIC DVS I (Branin test function)

DETERMINISTIC DVS II — Strong local search component (Branin test function)

- Convergence to an essential global minimizer of f with probability one can be demonstrated, provided f is lower semi-continuous and bounded below.

- **Definition** *An essential global minimizer x_* is a point for which the set*

$$L(x_*) = \{x \in \Omega : f(x) < f(x_*)\}$$

has Lebesgue measure zero.

CONVERGENCE (Deterministic)

- Convergence to an open set essential global minimum of f can be demonstrated, provided f is lower semi-continuous and bounded below.

- **Definition** *An open set essential global minimum $f_\#$ on Ω is the supremum of the values $f_0 \in \mathbb{R}$ for which the level set*

$$L(f_0) = \{x \in \Omega : f(x) < f_0\}$$

contains no open ball of positive radius.

COMPARISON WITH OTHER METHODS

	Problem	Ω	ARS	DIRECT	DVS I	DVS II
1	Ackley	$[-30, 30]$	2624	NA	1468	5027
2	Becker-Lago	$[-10, 10]$	118	81	516	149
3	Branin	$[-5, 10] \times [0, 15]$	153	107	639	140
4	Camel Back - 3	$[-5, 5]$	120	NA	254	115
5	Camel Back - 6	$[-5, 5]$	123	159	293	140
6	Cosine Mixture	$[-1, 1]$	192	NA	510	92
7	Dekkers-Aarts	$[-20, 20]$	1068	291	3805	1219
8	Goldstein-Price	$[-2, 2]$	191	123	1249	166
9	mod. Rosenbrock	$[-5, 5]$	3 fails	405	666	281
10	Schubert	$[-10, 10]$	329	2285	524	322
11	Beale	$[0, 5]$	214	329	687	434
12	Brown almost linear	$[-5, 25]$	226	197	751	269
13	Chebyquad	$[-1, 1]$	44	25	40	48
14	Freudenstein-Roth	$[-15, 15]$	2067	5619	6137	590
15	Jennrich-Sampson	$[-10, 10]$	18833	527	4440	4162
16	Trigonometric	$[-0.5, 1]$	102	85	209	63
17	Weka 1	$[0, 1]$	114	15387	37	30
18	Weka 2	$[0, 1]$	1722	1691	879	341
19	Weka 3	$[0, 1]$	4 fails	1861	48	27
20	offset Rastrigin	$[-1, 1]$	2929	749	1031	3228

Table : Number of function evaluations required to satisfy $f_k - f_* \leq$1e-3. ARS is stochastic so the results are averages over ten runs. Runs using more than 50000 function evaluations were halted and listed as 'fails'. Figures marked in red indicate the fastest method for each problem.

COMPARISON WITH OTHER METHODS

- Difficult problem:

$$f = \min\left(2 + x_1\cos(\theta) + x_2\sin(\theta), 40\|x - c\|_2\right) \quad \text{over} \quad [-1,1]^2,$$

where $\theta \in [0, 2\pi]$ and $c \in [-1,1]^2$ are chosen randomly.

- Ten different instances of this problem were generated. The average number of function evaluations to reduce f below 1e-3 are as follows:
  - DIRECT 12104
  - DVS I 7215
  - DVS II 4912
  - ARS failed on each run.

SUMMARY

- A dynamic partitioning algorithm for bound constrained global optimization called DVS has been presented.
- The method uses a series of Voronoi partitions to determine where the objective function should be evaluated.
- Provably convergent on nonsmooth problems (with probability one if random points are used).
- Numerical results show that DVS is effective and competitive on two-dimensional problems.
- The method becomes increasing inefficient in dimensions greater than two because the Voronoi cells become complicated.